

Exemplary haplotype diversity distribution (8 SNPs / 84Kb)

Potential SNP/Haplotype Combinations
 $2^{10} = 1024$

<u>Observed Haplotypes</u>	<u>Population Distribution (Frequency)</u>	
1. GGACAACC	283	(83.2%)
2. AATTCGGG	40	(11.8%)
3. GATTAGCC	2	(0.6%)
4. GGTCAGCC	2	(0.6%)

FIGURE 1

Exemplary Haplotype/SNP Allele State Matrix

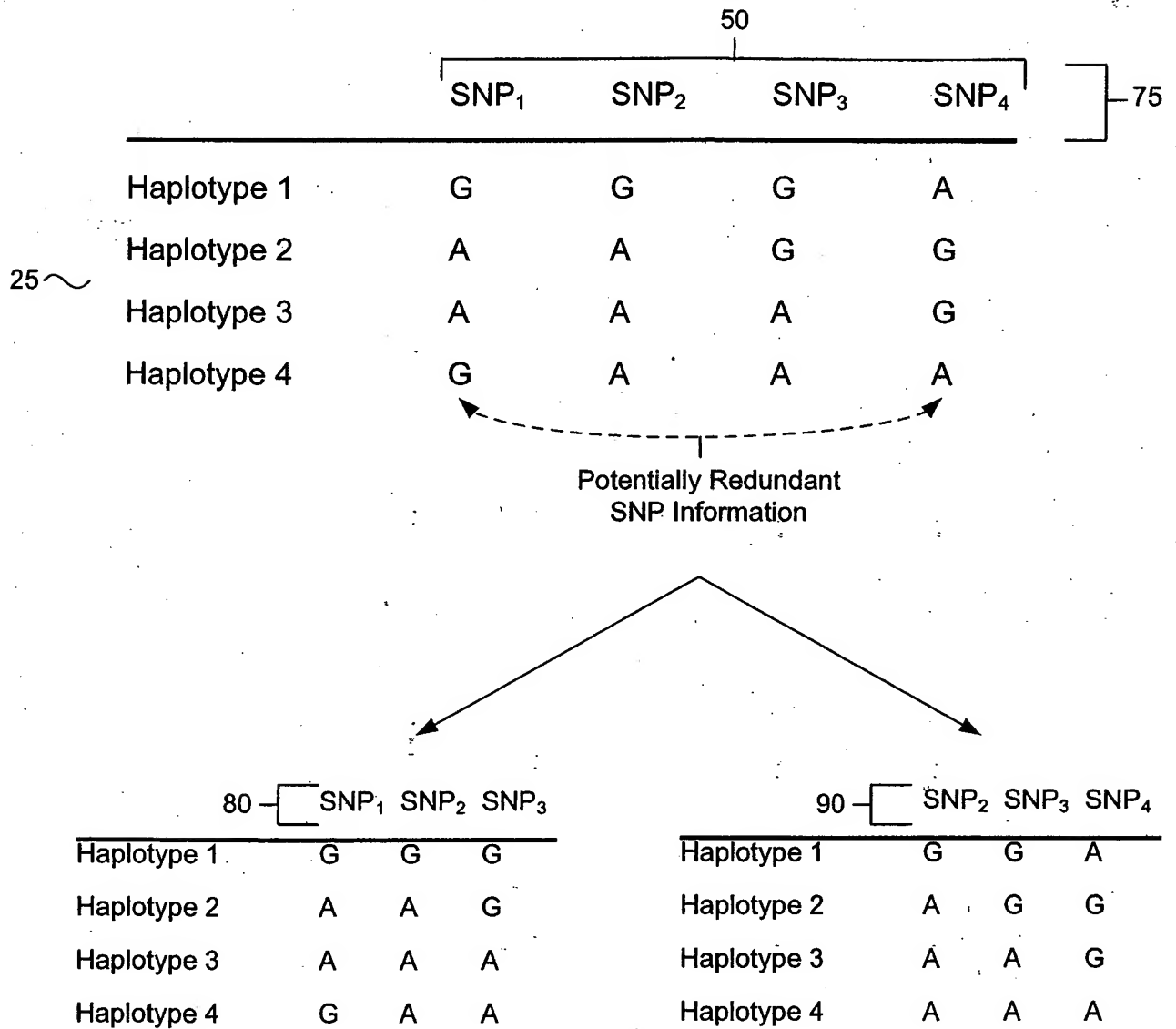


FIGURE 2

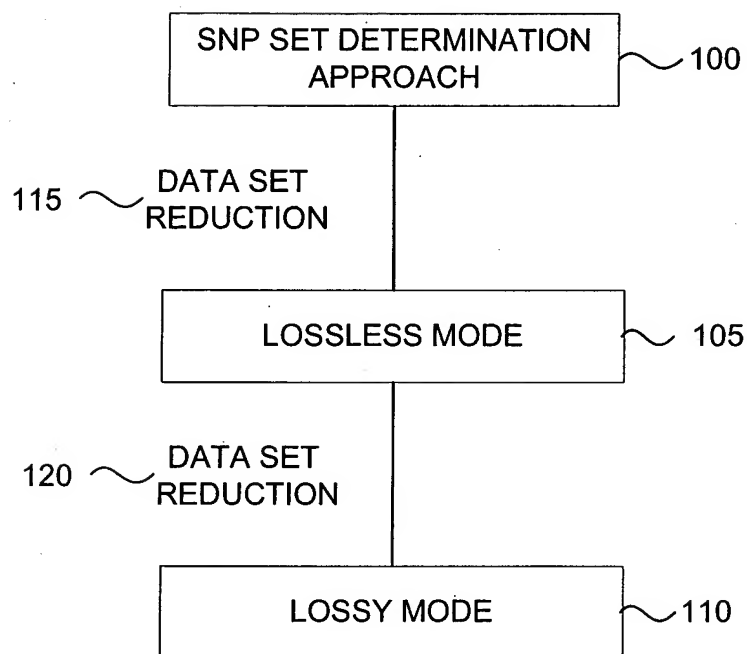


FIGURE 3

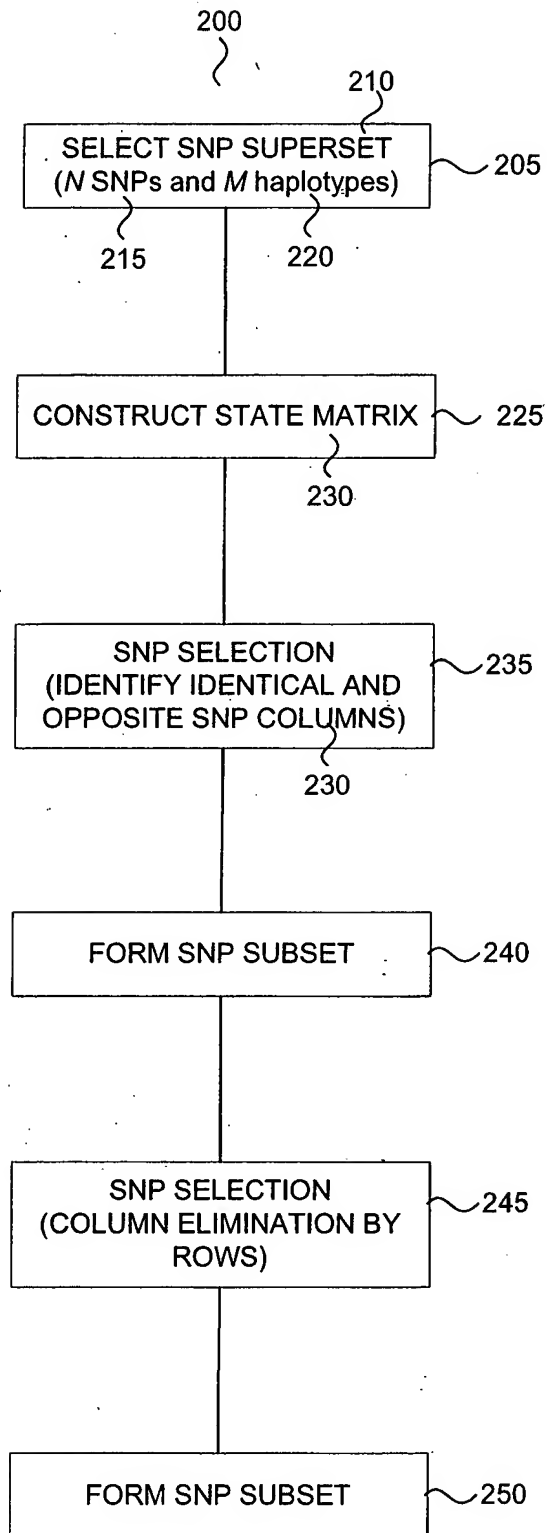


FIGURE 4A

221

Haplotype/SNP Allele State Matrix

224

	SNP ₁	SNP ₂	SNP ₃	SNP ₄	SNP ₅
Haplotype 1	G	G	G	A	G
Haplotype 2	A	A	G	G	A
Haplotype 3	A	A	A	G	A
Haplotype 4	G	A	A	A	A

222

FIGURE 4B

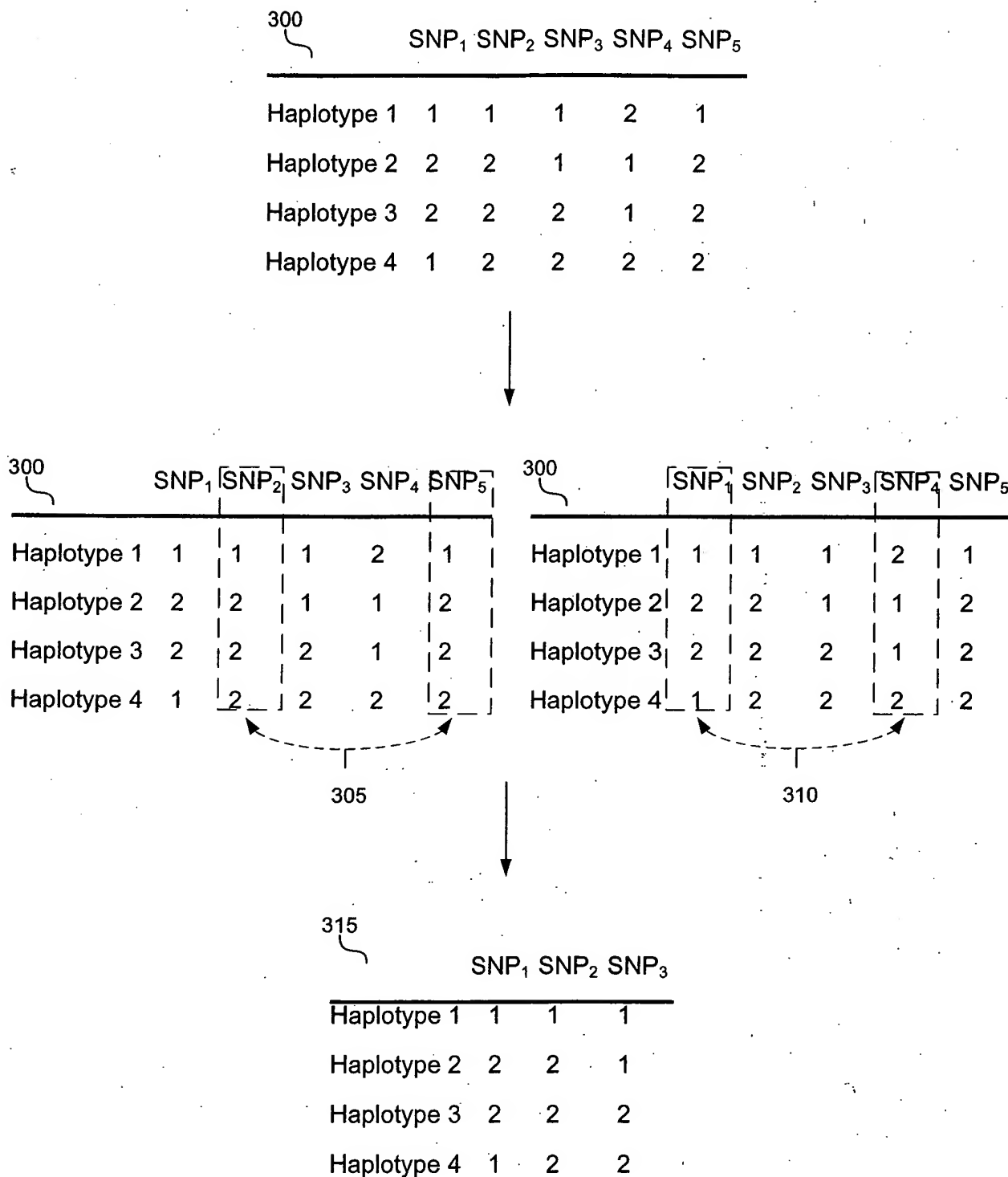


FIGURE 5A

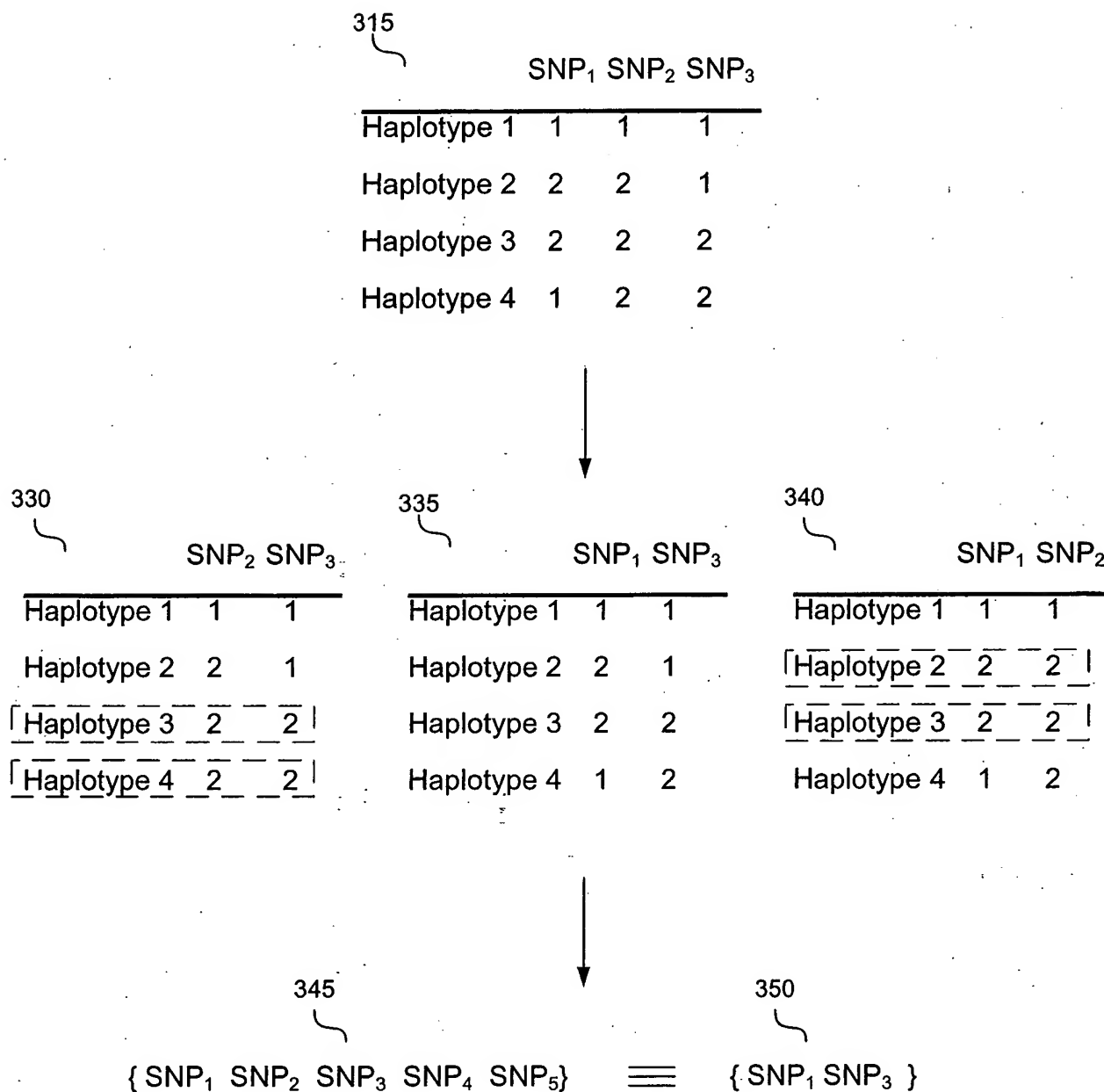


FIGURE 5B

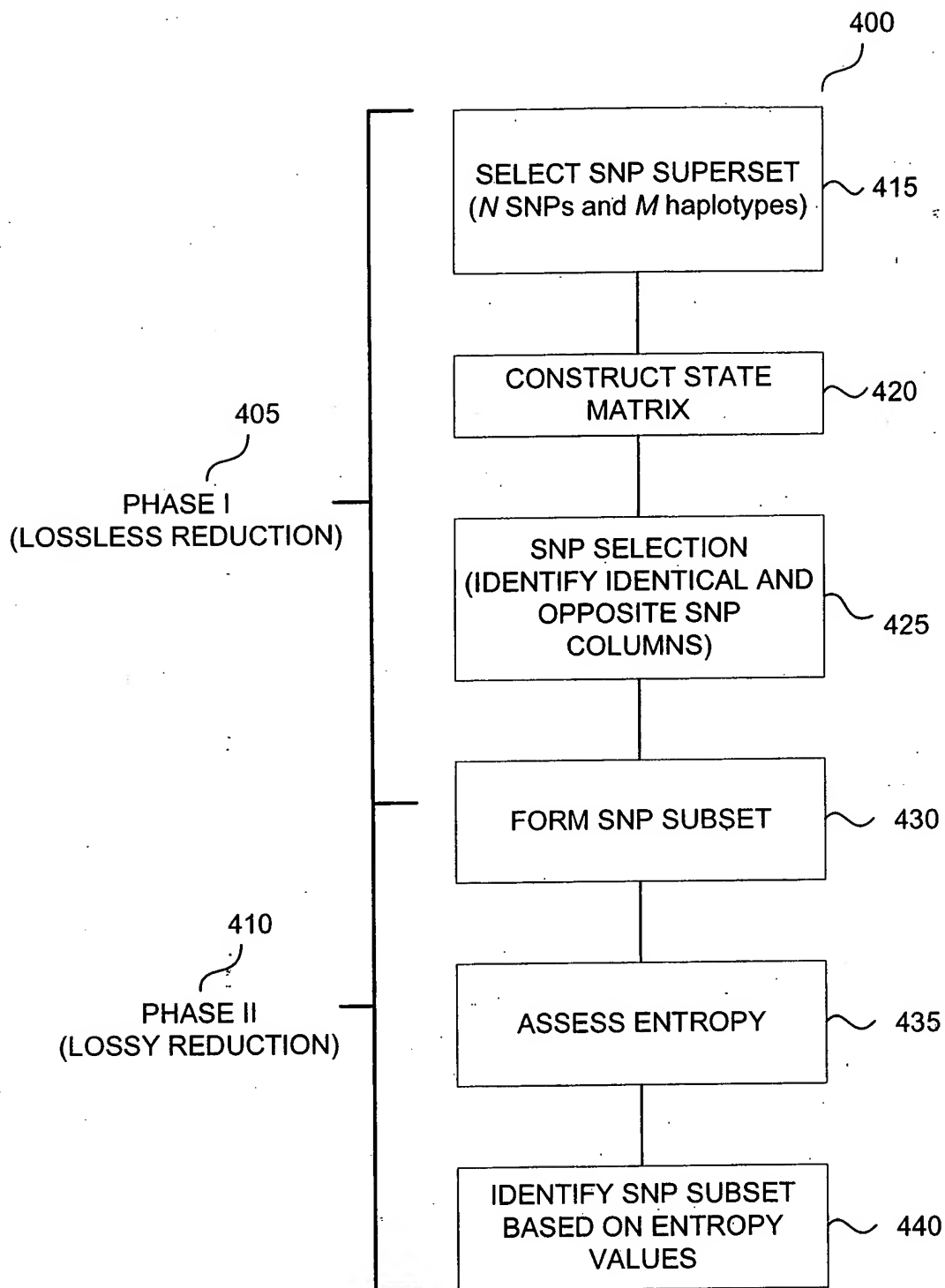


FIGURE 6A

ORIGINAL STATE MATRIX / SNP SUPERSET

455

460

454

M

"A"

N

Haplotype Number	P	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
1	0.1136	1	1	1	1	1	2	1	2	1	2	1	1	2	2	2	1	2
2	0.4318	1	1	1	1	1	2	1	2	1	2	1	1	2	2	2	2	2
3	0.0114	1	1	1	2	2	1	2	1	2	1	2	2	1	1	1	1	1
4	0.0454	1	2	2	1	1	2	1	2	1	2	1	1	2	2	2	1	1
5	0.0454	2	1	1	1	1	2	1	2	1	2	1	1	2	2	2	1	2
6	0.0118	2	2	2	2	2	1	2	1	2	1	2	1	1	1	1	1	1
7	0.3292	2	2	2	2	2	1	2	1	2	1	2	2	1	1	1	1	1
8	0.0114	2	2	2	2	2	1	2	1	2	2	1	2	1	1	1	1	1

FIGURE 6B

RESULTING STATE MATRIX FOLLOWING
APPLICATION OF LOSSLESS APPROACH

455

465

Haplotype Number	P	SNP ₁	SNP ₂	SNP ₄	SNP ₁₀	SNP ₁₂	SNP ₁₆	SNP ₁₇
1	0.1136	1	1	1	2	1	1	2
2	0.4318	1	1	1	2	1	2	2
3	0.0114	1	1	2	1	2	1	1
4	0.0454	1	2	1	2	1	1	1
5	0.0454	2	1	1	2	1	1	2
6	0.0118	2	2	2	1	1	1	1
7	0.3292	2	2	2	1	2	1	1
8	0.0114	2	2	2	2	2	1	1

FIGURE 6C

No. of SNPs (<i>k</i>)	No. of Combinations ***	Optimal Set of <i>k</i> SNPs	Haplotype Distribution Resulting from the Optimal SNP Set	Resulting Entropy (<i>H</i>) (bits)
7	8	{SNP ₁ , SNP ₂ , SNP ₄ , SNP ₁₀ , SNP ₁₂ , SNP ₁₆ , SNP ₁₇ }	(0.114, 0.432, 0.011, 0.045, 0.045, 0.012, 0.329, 0.011)	2.0351
6	28	{SNP ₁ , SNP ₄ , SNP ₁₀ , SNP ₁₂ , SNP ₁₆ , SNP ₁₇ }	(0.114, 0.432, 0.011, 0.045, 0.045, 0.012, 0.329, 0.011)	2.0351
5	56	{SNP ₁ , SNP ₁₀ , SNP ₁₂ , SNP ₁₆ , SNP ₁₇ }	(0.114, 0.432, 0.011, 0.045, 0.045, 0.012, 0.329, 0.011)	2.0351
4	70	{SNP ₁ , SNP ₁₂ , SNP ₁₆ , SNP ₁₇ }	(0.114, 0.432, 0.011, 0.045, 0.045, 0.012, 0.341)	1.9631
3	56	{SNP ₁ , SNP ₁₆ , SNP ₁₇ }	(0.114, 0.432, 0.057, 0.045, 0.352)	1.8475
2	28	{SNP ₁₂ , SNP ₁₆ }	(0.216, 0.432, 0.397)	1.5311
1	8	{SNP ₁₆ }	(0.5682, 0.4318)	0.9865

FIGURE 6D

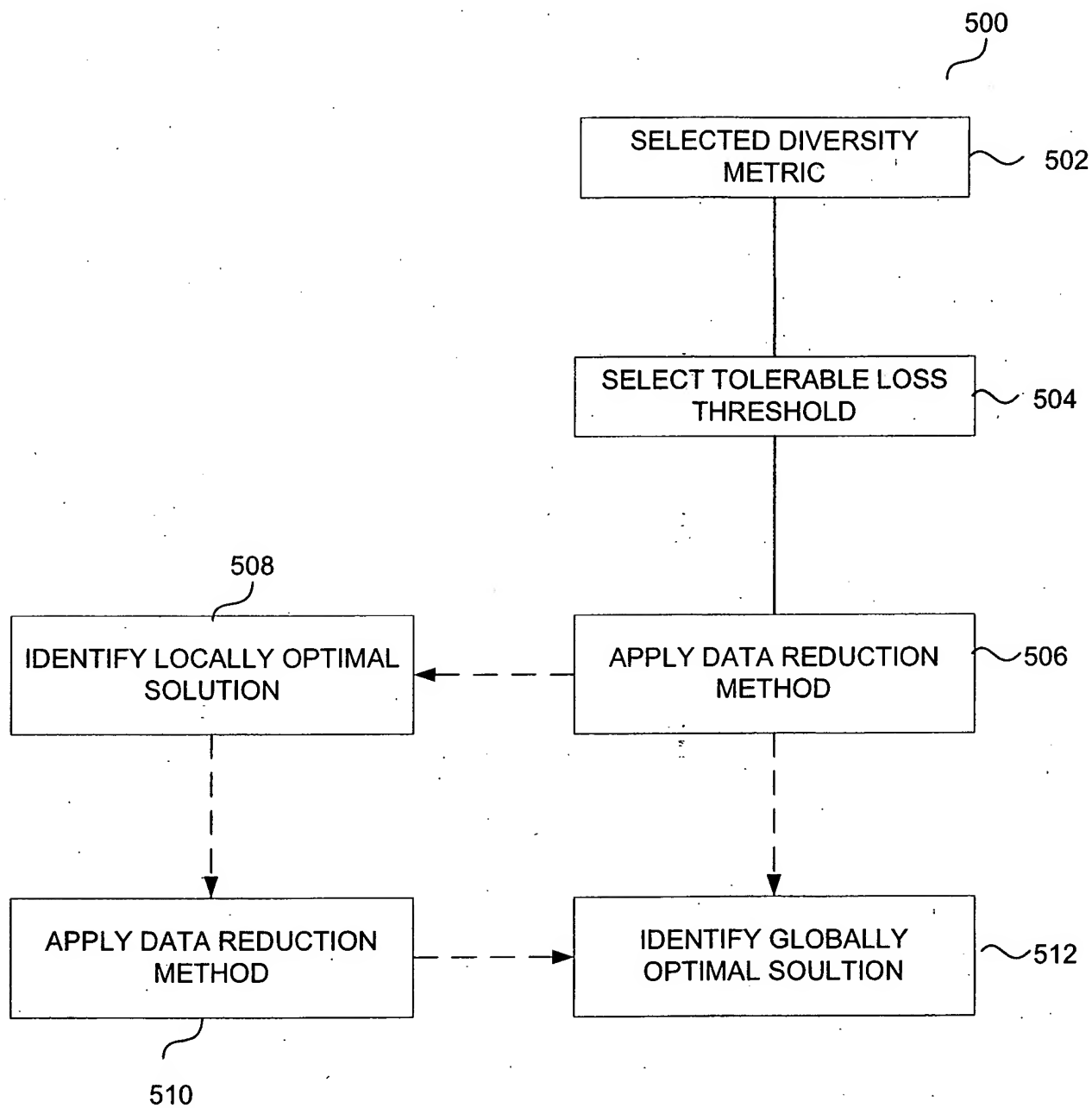


FIGURE 7

				557 ⌋	557 ⌋
	SNP ₁	SNP ₂	SNP ₃	SNP ₄	P
Hap1	1	1	1	2	0.63
Hap2	2	2	1	1	0.22
Hap3	2	2	2	1	0.07
Hap4	1	2	2	2	0.08

	No. of SNPs (K)	No. of Combinations $4! / (K! (4-K)!)$	Optimal Set of SNPs	558 ⌋ Remaining Info (H)
	4	1	{SNP ₁ SNP ₂ SNP ₃ SNP ₄ }	1.461
	3	4	{SNP ₁ SNP ₂ SNP ₃ }	1.461
559 ~	2	6	{SNP ₁ SNP ₃ }	1.461
	1	4	{SNP ₂ }	0.951

FIGURE 8

535

Chr.	Pop.	Total No. of SNPs	546 Mean Spacing Between SNPs (bp)	548 Mean Spacing Between SNPs in Genes (bp)	550 No. of Haplotype Blocks	552 Mean Block Size (bp)	556 Mean SNPs per Block	Mean Min. SNP per Block
6	A	2,504	24,386	10,840	646	23,000	3.88	Lossless 554
	C	4,009	23,694	10,630				
21	A	955	12,424	7,382	242	14,933	3.95	Loss 555
	C	1,555	11,921	7,031				
22	A	1,405	10,041	6,035	350	13,714	4.01	2.32
	C	1,783	9,080	7,760				
					417	17,505	4.28	2.27

540

545

FIGURE 9

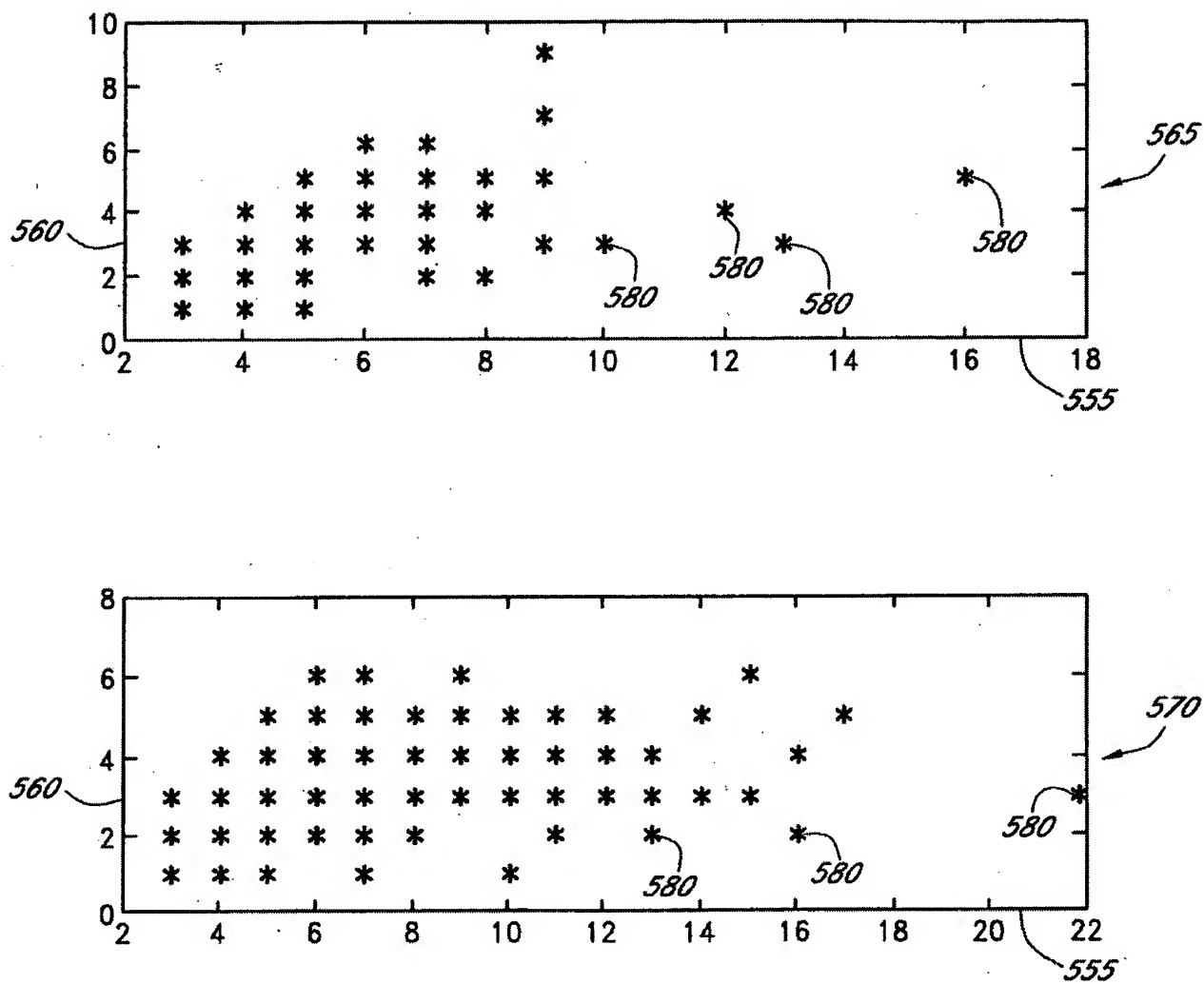


FIGURE 10